Email Spoofing Analysis

Nilay Mistry, Rishiraj Singh Bhati, Heena Jain, Mehul Parmar

Institute of Forensic Science

Gujarat Forensic Sciences University

Gandhinagar, India

***Abstract*—Email attacks are the one emerging now-a-days andare greatest threat to National Security. Many of the attackers are targeting high end audience and critical infrastructures. It is thus essential to identify and eliminate users and machines misusing e-mail service. There is a need to find a technique by which user can differentiate between the genuine and spoofed Email. As there are many ways to spoof an Email but this paper mainly focuses on mails spoofed by using either PHP Script or Anonymous Mailers. This paper represents techniques by which user can find whether the mail is spoofed or not.**

***Index Terms*—Email header, Message-Id, Email spoofing,Email Forensics, Email Header Analysis, PHP Mailer, Anonymous Mailer.**

INTRODUCTION

Email consists of 3 parts:

* Envelope contains the internal process by which Email is routed.
* Body contains the actual contents of the message.
* Header contains the complete information about the Email like From, Date, Message-Id, To, Subject, Received, and Sender.

It uses RFC (Request for Comment) 6532[1] as a standard for creation of Email header and RFC 5322[2] for Email body. Email is the most used means of communication between people for their day to day operations in a private company or government organizations. Now a day’s many attackers are targeting high end audience and critical infrastructures that can harm the national security. Even an insignificant Email can harm the whole national security. So, spoofed Email can be considered as the smallest weapon of destruction that can create a big chaos.

Email spoofing is a mechanism in which Email appears to come from authentic user but actually it comes from an imposter. Email Spoofing can involve not only sender email id but also MIME, Message-Id[4], or any activity that avoids tracing of Email. Email spoofing is very simple to do because the core protocol does not provide any authentication. Even a non technical person can do so by using many anonymous mailers that are freely available on the internet.

Earlier FTP (File Transfer Protocol) and HTML (Hyper Text Markup Language) was used to exchange Email over the ARPANET but now it uses SMTP (Simple Mail Transfer Protocol) and ESMTP (Extended SMTP) as an internet standard

for its transmission. To spoof an Email, attacker forges the SMTP by changing the attributes of Email header according to his/her need. In this some of the fields can be changed but not all. The fields like From, To, Bcc, Cc etc can be changed, there are likely hood of some of the attributes like Message-Id, SPF (Sender Policy Framework), DKIM[5] (Domain Keys Identified Mail), and DMARC (Domain-based Message Authentication, Reporting and Conformance) are also subjected to change. These fields can be used to identify whether an Email is genuine or spoofed. Further IPv6 regime offers more security on Email spoofing over IPv4.

Email service providers like Gmail, Yahoo and Windows Live uses special algorithm to filter out the spoofed Emails but then even there are chances that spoofed mail appears in Inbox. But some of the private companies and government organizations that use their own Email servers for communication unlike Google, Yahoo and Windows Live does not have any mechanism that filters out the spoofed Email and send it to spam folder. Thus to identify this, an algorithm can be implemented that provides and extra layer of security to the Email servers.

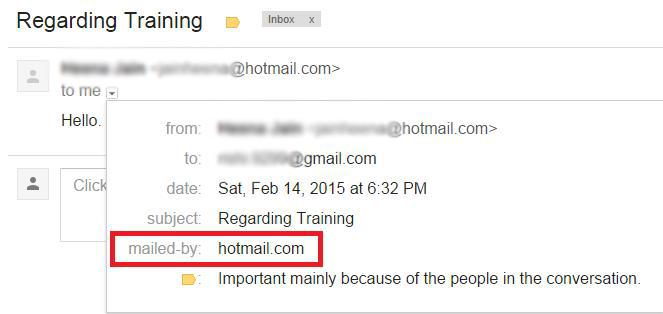


Fig. 1. Genuine Email

This mail is sent by genuine hotmail account.

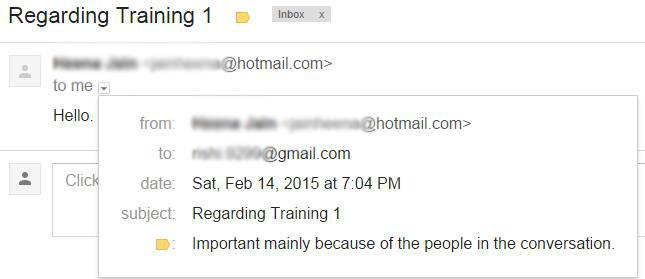


Fig. 2. Spoofed Email

This mail is sent via a PHP script using open source PHP Mailer as a library. As stated it is a spoofed mail so it should be received in spam or junk but it is received in inbox.

Both mails look completely similar but original mail has an extra attribute i.e. mailed-by which distinguishes it from spoofed.

To check which servers are vulnerable to this kind of attack both genuine and spoofed Emails are sent and checked whether they are received in Inbox or Spam. Later on all the Email headers are gathered and checked what the differences are there between original and spoofed Emails.Now let’s see how to identify the spoofed Emails by analyzing there headers.

ANALYSIS

In this complete analysis of Email headers [6] is done by comparing different kinds of Emails sent on different Email servers using different ways, both genuine and spoofed. Email service providers used for analysis are:

* Gmail
* Yahoo
* Windows Live
* Mail.India.com

Three type of Email are analyzed:

* Genuine Email
* Spoofed Email using 3rd party anonymous mailer (www.emkei.cz).
* Spoofed Email using our own script written in PHP. Now all the possible combinations of these Email service

providers with different ways of sending Email are made and the Emails are sent. Some of the Emails were received but some of them were not. Some spoofed Emails look like genuine but some were flagged as spam. This table represents how Email service providers respond to spoofed Emails.

Table I. Response to Spoofed Emails

|  |  |  |
| --- | --- | --- |
| **Source** | **Destination** | **Response** |
|  |  |  |
| Gmail | Gmail | Received in Inbox but |
|  |  | flagged as spam |
| Gmail | Yahoo | Received in Inbox |
| Gmail | Windows Live | Received in Spam |
| Gmail, Yahoo, | India.com | Received in Inbox |
| Live |  |  |
| Yahoo | Yahoo | Not Received |
| Yahoo | Gmail | Not Received |
| Yahoo | Windows Live | Not Received |
| Hotmail | Windows Live | Received in Inbox |
| Windows Live | Gmail | Received in Inbox |
| India.com | India.com | Received in Inbox |
| India.com | Gmail | Received in Inbox |

After sending the Emails their Email headers were retrieved and then completely analyzed to understand the behavior of spoofed Emails. Below are all the comparison tables that were made in accordance to understand the differences. It is also

mentioned that the selected framework is based on popularity and the study has been restricted to the mentioned attributes.

Table II. Gmail to Others

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Attributes** | **Original** | **Spoofed** | **Spoofed** |  |
| **Emkei.cz** | **PHP Script** |  |
|  |  |  |
| SPF | Pass | Softfail | Softfail |  |
| Received-SPF | No such word | Domain of | Domain |  |
|  |  | transitionin | of |  |
|  |  | g <Email> | transitioni |  |
|  |  | does not | ng |  |
|  |  | designate | <Email> |  |
|  |  | <ip> as | does not |  |
|  |  | permitted | designate |  |
|  |  | sender | <ip> as |  |
|  |  |  | permitted |  |
|  |  |  | sender |  |
| Message-Id | ID@mail.gmai | ID@emkei. | ID@www |  |
|  | l.com | cz | .drooper.i |  |
|  | ID is of 51 | ID is less | n |  |
|  | characters | than 45 char | ID is less |  |
|  |  |  | than 45 |  |
|  |  |  | char |  |
| Dkim | Available/Pass | Neutral | Neutral |  |
| Dmarc | NA | Fail | Fail |  |
| x-hmca | Pass | Fail | Fail |  |
| x-auth-results | Pass | Fail | Fail |  |
| x-sid-results | Pass | Fail | Fail |  |
| To | Name | Email | Email |  |
|  | <Email> |  |  |  |
| From | Name<Email> | “Name”<E | Email |  |
|  |  | mail> |  |  |
| X-AntiAbuse | NA | NA | Displays |  |
|  |  |  | original |  |
|  |  |  | Email id |  |
|  |  |  | used for |  |
|  |  |  | authentica |  |
|  |  |  | ting |  |
| X-Priority | NA | Available | NA |  |
| X-Spam-Score | -1.899 | 0.8 | 1.1 |  |
| X-Spam-Level | NA | NA | \* |  |
| X-Spam- | No | No | No |  |
| Status |  | (But it | (But it |  |
|  |  | should be | should be |  |
|  |  | Yes) | Yes) |  |

Table III. Windows Live to Others

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Attributes** | **Original** | **Spoofed** | **Spoofed** |  |
| **Emkei.cz** | **PHP Script** |  |
|  |  |  |
| Spf | Pass | Softfail | Softfail |  |
| xhmca | Pass | Fail | Fail |  |
| x-auth-results | Pass | Fail | Fail |  |
| x-sid-results | Pass | Fail | Fail |  |
| Received-SPF | No such word | Domain of | Domain |  |
|  |  | transitionin | of |  |
|  |  | g <Email> | transitioni |  |
|  |  | does not | ng |  |
|  |  | designate | <Email> |  |
|  |  | <ip> as | does not |  |
|  |  | permitted | designate |  |
|  |  | sender | <ip> as |  |
|  |  |  | permitted |  |
|  |  |  | sender |  |
| X-TMN | Available | NA | NA |  |
| Dkim | Available | NA | NA |  |
| Dmarc | NA | Fail | Fail |  |
| To | Name | Email | Email |  |
|  | <Email> |  |  |  |
| X-AntiAbuse | NA | NA | Displays |  |
|  |  |  | original |  |
|  |  |  | Email id |  |
|  |  |  | used for |  |
|  |  |  | authentica |  |
|  |  |  | ting |  |
| Message-Id | ID@phx.gbl | ID@emkei. | ID@www |  |
|  | ID is of 35 | cz | .drooper.i |  |
|  | characters | ID is less | n |  |
|  |  | than 25 char | ID is less |  |
|  |  |  | than 32 |  |
|  |  |  | char |  |
| From | Name<Email> | “Name”<E | Email |  |
|  |  | mail> |  |  |
| Received | Name of | Name of | Name of |  |
|  | original server | original | original |  |
|  |  | server | server |  |
| X-spam-Flag | Yes | No | No |  |
| X-Spam-Score | 4.001 | 2 | 1.501 |  |
| X-Spam-Level | \*\*\*\* | \*\* | \* |  |
| X-Spam- | NO | No | No |  |
| Status |  | (But it | (But it |  |
|  |  | should be | should be |  |
|  |  | Yes) | Yes) |  |

Table IV. Yahoo to Others

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Attributes** | **Original** | **Spoofed** | **Spoofed** |  |
| **Emkei.cz** | **PHP Script** |  |
|  |  |  |
| X-spam-Flag | Yes | No | No |  |
| X-Spam-Score | 10.724 | 3.13 | 2.231 |  |
| X-Spam-Level | \*\*\*\*\*\*\*\*\*\* | \*\*\* | \*\* |  |
| X-Spam- | Yes | No | No |  |
| Status | (But it should | (But it | (But it |  |
|  | be NO) | should be | should be |  |
|  |  | Yes) | Yes) |  |
| Received | Name of | Name of | Name of |  |
|  | original server | original | original |  |
|  |  | server | server |  |
| X-AntiAbuse | NA | NA | Displays |  |
|  |  |  | original |  |
|  |  |  | Email id |  |
|  |  |  | used for |  |
|  |  |  | authentica |  |
|  |  |  | ting |  |
| Message-Id | ID.JavaMail.y | ID@emkei. | ID@www |  |
|  | ahoo@mail.ya | cz | .drooper.i |  |
|  | hoo.com |  | n |  |
| From | Name<Email> | “Name”<E | Email |  |
|  |  | mail> |  |  |
|  | Table V. India.com to Others | |  |  |
|  |  |  |  |  |
| **Attributes** | **Original** | **Spoofed** | **Spoofed** |  |
| **Emkei.cz** | **PHP Script** |  |
|  |  |  |
| X-Spam-Score | 4.723 | 1.5 | 1.1 |  |
| X-Spam-Level | \*\*\*\* | \* | \* |  |
| X-Spam- | No | No | No |  |
| Status |  | (But it | (But it |  |
|  |  | should be | should be |  |
|  |  | Yes) | Yes) |  |
| Received | NA | Name of | Name of |  |
|  |  | originating | originatin |  |
|  |  | server | g server |  |
| X-AntiAbuse | NA | NA | Displays |  |
|  |  |  | original |  |
|  |  |  | Email id |  |
|  |  |  | used for |  |
|  |  |  | authentica |  |
|  |  |  | ting |  |
| Message-Id | ID.JavaMail.to | ID@emkei. | ID@www |  |
|  | mcat@be04 | cz | .drooper.i |  |
|  |  |  | n |  |
| From | Name<Email> | “Name”<E | Email |  |
|  |  | mail> |  |  |
| X-Mailer | zMail Mailer | NA | PHPMaile |  |
|  | (beta) |  | r |  |

INFERENCE

The attributes that is necessary to identify whether an Email

is spoofed or not were used in the proposed framework and is

also shown below.

The attributes selected for the study which is also part of the framework selected are assumed to be less likely of being spoofed. The selected attributes play a major role in providing authenticity, confidentiality and integrity while studying the scope of spoofing. Fig 3 shows the attributes and its description. Subsequently it is inference that among all the email services that were included in this study, it was found that certain emails were less likely of being spoofed than the other.

Spf

Message

Id

Dkim

Signature

X-

Antiabus

Received

Authenti

cated Id

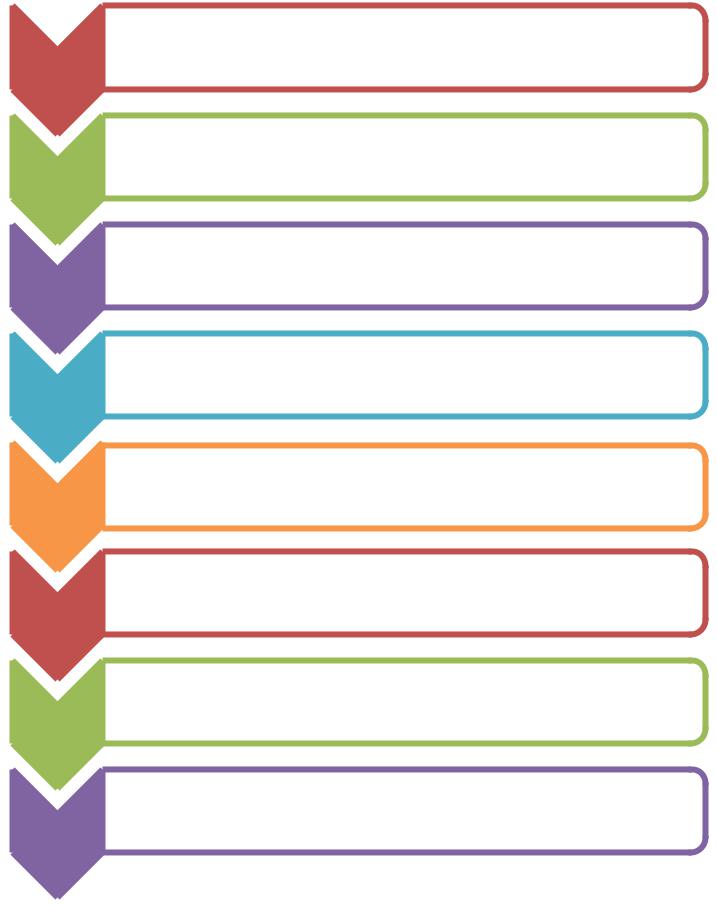
X-Spam

Score

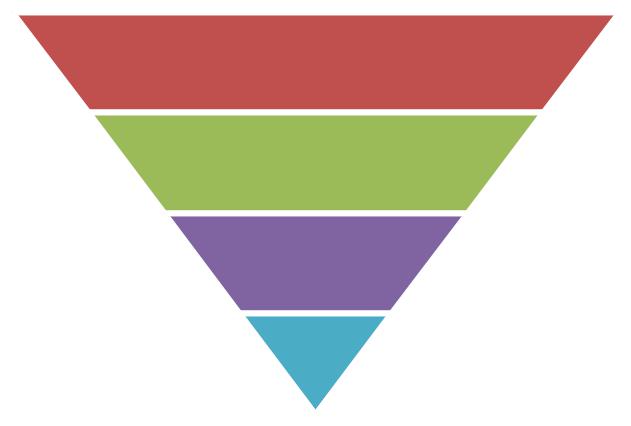
X-Spam

Status

* Sender Policy Framework is used to authenticate the sender and its value must be "PASS"
* It is used to uniquely identify the Email and it is the combination of timestamp,process id,user id followed by @original domain.
* Domain Keys Identified Mail is used to provide authentication Framework at Domain Level and to protect message signer identity and integrity of message.
* It is present in Spoofed mail which can be used to track abuse as it has Primary Hostname , Original Domain , Originator/Caller UID/GID , Sender Address Domain .
* It is used to trace the email to origin as it shows the details of the route from were mail has passed.
* It is present in X-Antiabuse which shows the original mail id from where the mail is sent.
* It shows the score which let us know about the level of Spam.
* It tell us wheather the mail recieved is spam or not .



**India.com**



**Windows Live**

**Gmail**

Yahoo

Fig. 3. Attributes to Identify Spoofed Email

Fig. 4. Likely hood of spoofing of Email header

CONCLUSION AND FUTURE WORK

This paper was completely based on the analysis of Email header to identify whether it is spoofed or not. Here, around 50 Email headers have been compared with one other using all the important attributes that define the originality and source of the Email. In future this analysis can be used to build an Email server that is intelligent enough to identify whether the Email is spoofed or not. It will check each and every attribute that is essential to consider the authenticity of an Email. Today all the government Email servers are most prone to Email headers spoofing and even an insignificant Email can create a big chaos. So to protect confidentiality an extra layer of security is compulsory for Email servers.

ACKNOWLEDGMENT

“It is not the brain that matter the most, but that which guide them: The character, the heart, generous qualities and progressive force.” We would like to make a number of acknowledgements to those who have helped us to prepare this paper. The divine support given by our guide Mr. Nilay Mistry (Assistant Professor, Institute of Forensic Science, GFSU, Gandhinagar), without whom the work would not possible.

We are highly grateful to Dr. M.S. Dahiya (Director, IFS, GFSU, Gandhinagar) for providing us this opportunity to carry

out independent study on this topic. Finally we offer our sincere thanks to our institute for their unflagging efforts for bringing out this paper.

REFERENCES

1. A.Yang, ”Internationalized Email Header” Internet Engineering Task Force (IETF), 2012 RFC 6532.
2. P. Resnick, “Internet Message Format” Internet Engineering Task Force (IETF), 2008 RFC 5322.
3. P. Resnick, “Internet Message Format” Internet Engineering Task Force (IETF), 2001 RFC 2822.
4. Satheesaan Pasupatheeswaran, “Email 'Message-IDs' helpful for forensic analysis?” Originally published in the Proceedings of the 6th Australian Digital Forensics Conference, Edith Cowan University, Perth Western Australia, December 3rd 2008.



1. Allma E., Callas J., Delan M., Libbey M., Fenton J. & Thomas

M., “Domain Keys Identified Mail (DKIM)”, Internet

Engineering Task Force (IETF), 2007, RFC 4871.

1. M. Tariq Banday, “Analysing E-Mail Headers for Forensic Investigation”, Journal of Digital Forensics, Security and Law, Vol. 6(2).
2. B. Radvanovsky, “Analyzing spoofed e-mail header” Journal of Digital Forensic Practice, 1(3), 2006, pp. 231-243.
3. Landsiedel O., Niedermayer H., Wehrle K, “An Infrastructure for Anonymous Internet Service In IWI2005, Chiba/Tokyo, Japan, May 2005.
4. Crocker, D. (2009), “Internet Mail Architecture”, RFC 5598,

July 2009.http://tools.ietf.org/pdf/rfc5598.pdf, (25-Mar-2011).

1. Ehrenkranz, T. and Li, J. On the state of IP spoofing defence. In Proceedings of ACM Trans. Internet Technician 2009.

**Less**

**More**